Washington State Wine Commission
Demonstration Research Grant Program

Final Report for FY 2020-21

Deadline for Submission: **June 30, 2021**

Date of Submission: **June 30, 2021**

Submit Final Report to Washington State Wine Commission: mhansen@washingtonwine.org

PROJECT TITLE: Crop Estimation Through Remote Sensing

Project Duration: 1 year

Principal Investigator: William Beightol
Organization: Collab Wine Company
Telephone: 509.832.4191
Email: will@collabwine.com

Co-PI: Ben Utter
Organization: Bloomfield Robotics
Telephone: 781.296.9929
Email: ben@bloomfield.ai

Cooperators:

Grower cooperators - Mercer Ranches, Phinny Hill Vineyards, Ste. Michelle Wine Estates, Winemakers LLC, Coyote Canyon Vineyards, Shaw Vineyard, McKinnly Springs Vineyard

Project cooperator - Dr. Wade Wolfe

Objectives:

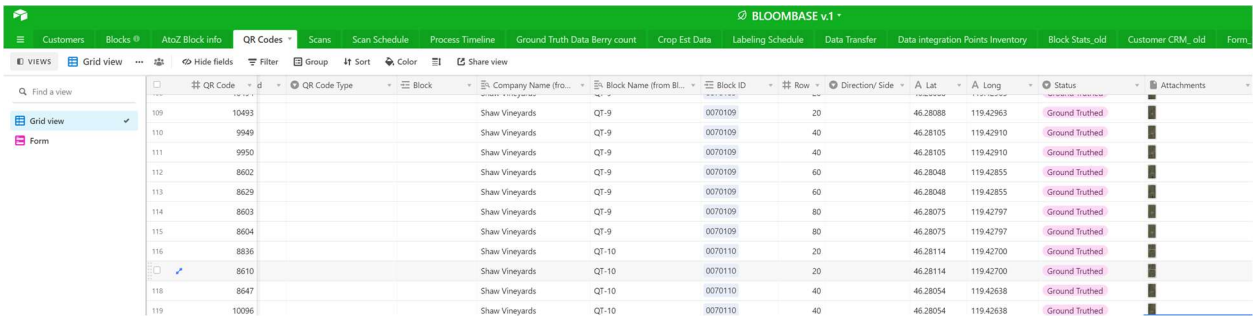
To provide growers actionable information on the crop load of their vineyard blocks near lag phase and véraison through beta-testing the use of Bloomfield Robotics "Flash" technology and analysis through their AI (Artificial Intelligence) system.

Additionally, to compare crop estimation results with other methods being trialed in other Washington State Wine Commission projects.

Field Operations:

Field operations as a whole during this project was a collaborative effort between Collab Wine and Bloomfield, as we combined knowledge bases and developed practices around scan scheduling and data tracking. These scans included scanning 2 blocks at each of 7 grower partner blocks at least two times each and at different timing around lag phase, véraison, and pre-harvest. This resulted in 36 different block scans. As scans progressed throughout the season, it became clear that a customer and scan database was an essential part of field operations and the collaboration. Collab Wine introduced the AirTable database development platform (pictured below) and created a new relational database that has evolved into the central hub for Bloomfield's operational data for customers,

blocks, scans, and extensive metadata that is tracked for each.



In addition to creating the foundation for our continued collaborative field operations, Bloomfield's team was educated on essential viticulture concepts and vineyard management practices and exited the 2020 season with a firm grasp on growth phases, block layouts, vine training systems, and grower needs particularly focused on Washington vineyards.

Field Scanning:

ATV driving patterns were quickly developed to enable efficient scanning of the morning/eastern side of the vines based on targeted coverage percentages. The percentage of coverage was generally aimed at 25% of the rows scanning the morning side of the canopy. Targeting the morning side usually allowed for the clearest view of the fruit zone, as this is the area most growers target for leaf removal. The variation in leaf removal by the different vineyard managers along with the time that has elapsed since leaf removal creates more variables to be accounted for in the crop estimation formula. Camera mounting was made more compact and efficient through the design and manufacturing of a steel plate that can be easily bolted onto common ATV front racks (pictured below). The field scanning was demonstrated with our project cooperator, Dr Wade Wolfe, during a lag phase timed scan at Andrews Vineyard on July 10th, 2020.

Ground Truthing:

Ground truthing methods evolved to combine traditional vineyard scouting techniques with Bloomfield's need to designate the ground truth (calibration) vines with markers that can be easily read by computer vision. Weather-proof QR code markers were attached to trellis wires and posts. These QR codes are then detected during the Flash log extraction process, and AI results for these plots are compared against human ground truthing. The results are then used to further calibrate yield predictions, adjusting for clusters that are partially or completely occluded and thus not detectable from imagery.

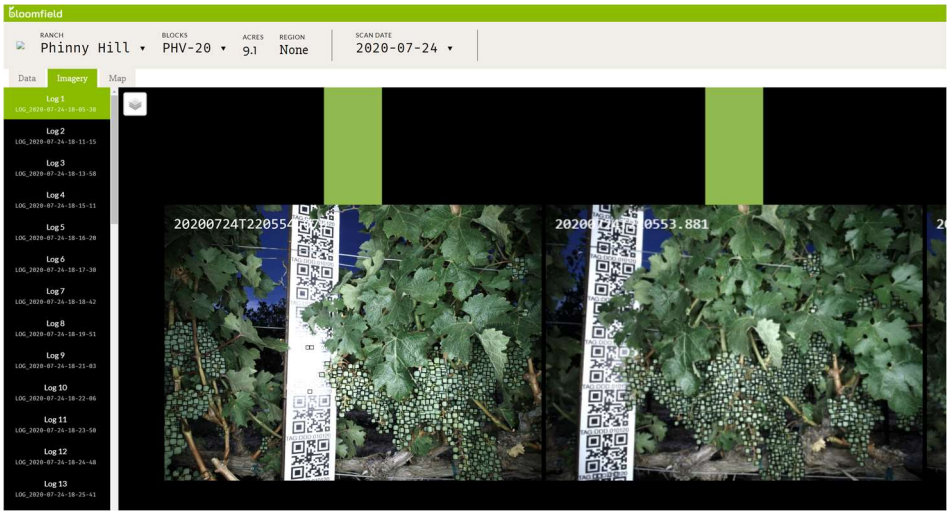
The ground truthing has many aspects that are used in traditional grape crop estimation including counting clusters, destructive harvest of clusters for weighing, and berry sampling to help determine average berry weight in an effort to determine an estimation of the number of berries per cluster.

Data Pipeline:

Flash generates a large amount of data and transferring it into the cloud for processing through various computer vision and AI analyses presented a challenge. A co-location arrangement was established with a local internet service provider to provide a very high-speed internet connection for upload to the cloud. Data is first transferred off Flash onto a hard drive, and the hard drive is then delivered to the co-location provider for insertion into a Bloomfield-managed server. Data upload generally completes within 24 hours. While we need to further improve and accelerate this process, this arrangement enables upload within a commercially viable timeframe.

Once the data was uploaded to AWS for analysis, we ran into significant delays in getting the data extracted and delays in models build that would allow for Data Analysis. These delays did not allow for an actionable insight to produce during the growing season. As of the publishing of this report, we are still waiting for insights that could serve inputs into a harvest estimation formula.

In the image below, you can see the Berry Detection model being run on a post-Lag Phase scan preformed at Phinny Hill Vineyards on July 24th.



This also gives you a view into how the QR Codes are used to set up the ground truthing sections similar to a “reference” vine setup used by some viticulture teams.

Data Analysis:

The extraction and analysis of the imagery generated by Flash and transformation into usable insights is an area of significant complexity. Over the course of the period partially supported by this grant, the extraction and analysis processes evolved from a collection of scripts inherited from Carnegie Mellon research into a commercial quality pipeline that is growing more scalable by the day.

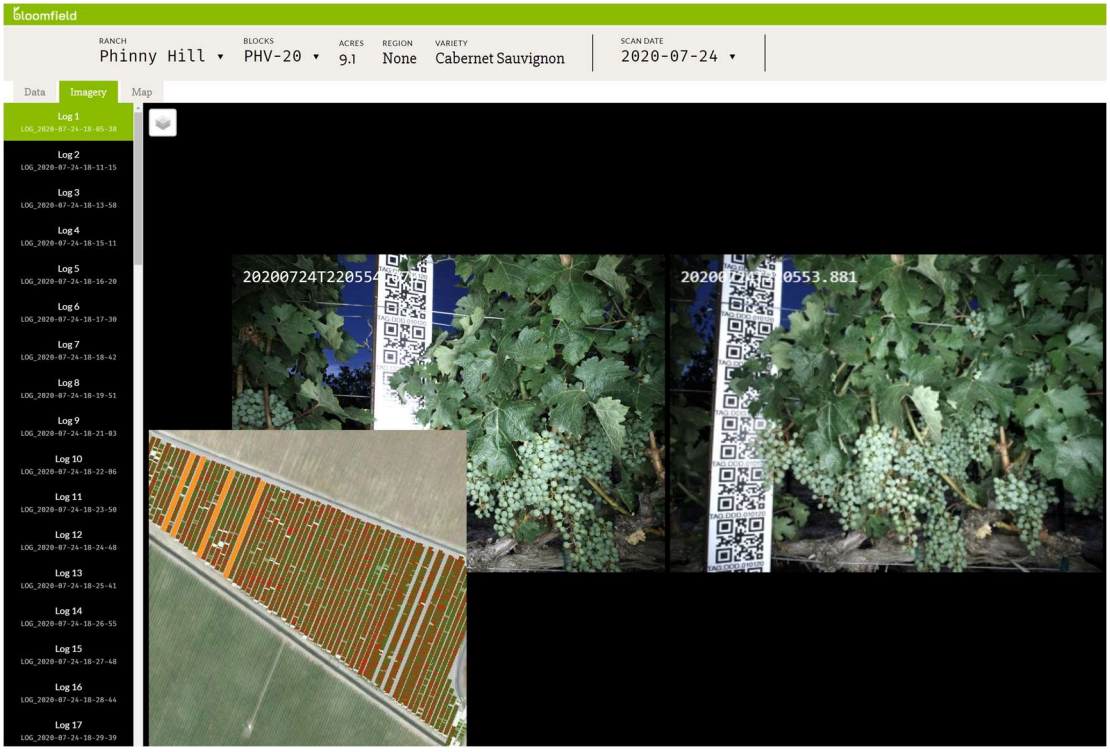
AI development practices (increasingly referred to as “MLops” or Machine Learning Operations) at Bloomfield have evolved to enable much more rapid human annotation (the training data for the AI), model development, and ongoing model training and tuning.

Producing highly accurate yield estimation is still an active area of development from both Bloomfield data scientists and processing pipeline specialists. Yield estimation is going to take considerable amount of time to be commercially viable and still contends with many of the factors that face viticulturist today, including a growth factor from the date of the scan until harvest. One of the most promising theses is that year over year scanning will inform the grower on how the crop is comparative to the last vintage. We are continuing to scan a few of the blocks that were included in the study to further test this thesis.

The delivery of the information back to the growers and viticulturist is accomplished through the Bloomfield.ai dashboard. The dashboard has been through a couple of development cycles during the time of this project with the newest version (pictured below - top) seeing a big improvement of the second quarter of 2021. The images below are from Alder Ridge Vineyards Lag Phase Scan along with a Verasion scan preformed at McKinley Springs Vineyard.

The latest models and outputs being run against last year’s scans are starting to produce some the desired results including a heat map of berry detections across an entire block. One of the desired outcomes is the location of missing vines along with a good picture on where thinning resources could be committed in a block to quickly even out or reduce congestion (below-bottom).


bloomfield
winemakers-llc (AR)
RANCH
Alder Ridge Vineyard
BLOCKS
AR-49
ACRES
15
REGION
None
SCAN DATE
2020-07-07
Data
Imagery
Map


bloomfield
RANCH
Phinny Hill
BLOCKS
PHV-20
ACRES
9.1
REGION
None
VARIETY
Cabernet Sauvignon
SCAN DATE
2020-07-24
Data
Imagery
Map
Log 1
LOG_2020-07-24-18-05-38
Log 2
LOG_2020-07-24-18-11-15
Log 3
LOG_2020-07-24-18-13-58
Log 4
LOG_2020-07-24-18-15-11
Log 5
LOG_2020-07-24-18-16-20
Log 6
LOG_2020-07-24-18-17-30
Log 7
LOG_2020-07-24-18-18-42
Log 8
LOG_2020-07-24-18-19-51
Log 9
LOG_2020-07-24-18-21-03
Log 10
LOG_2020-07-24-18-22-06
Log 11
LOG_2020-07-24-18-23-50
Log 12
LOG_2020-07-24-18-24-48
Log 13
LOG_2020-07-24-18-25-41
Log 14
LOG_2020-07-24-18-26-55
Log 15
LOG_2020-07-24-18-27-48
Log 16
LOG_2020-07-24-18-28-44
Log 17
LOG_2020-07-24-18-29-39

Data Comparison:

The comparison of data with Manoj Karkee was planned to be conducted on Phinny Hill Vineyards block 20. I met with Dr Karkee and his interns so that they could gain access to the block and use it in their yield estimation project. Unfortunately, from the Bloomfield side, we were unable to produce a yield estimation for comparison.

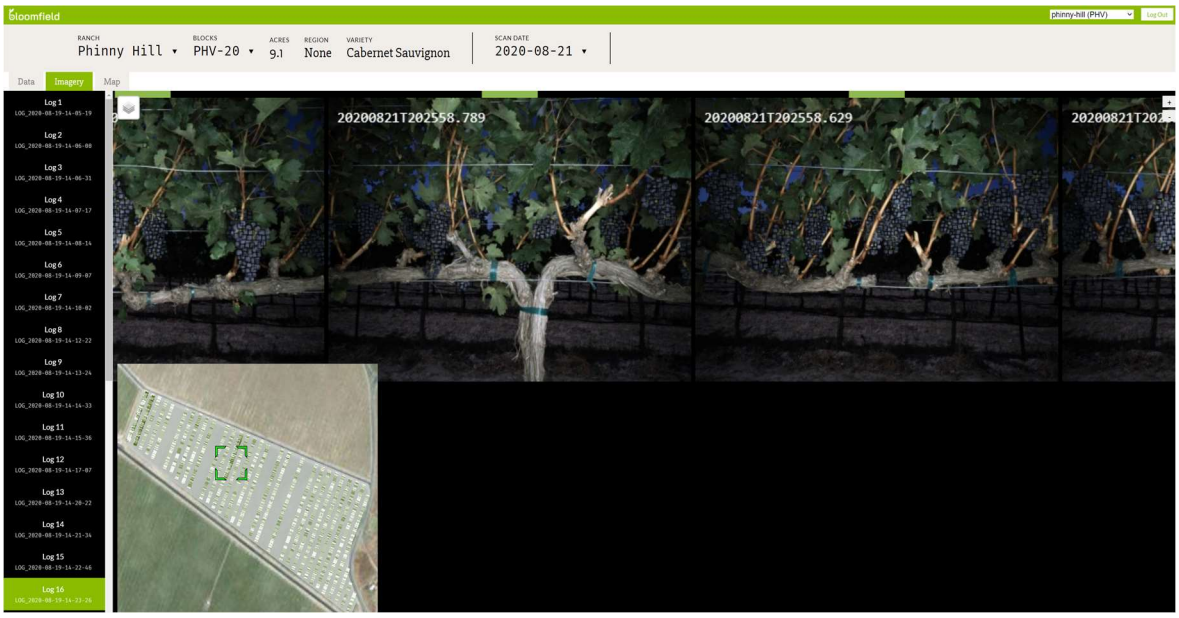
Data Availability:

The goal from the beginning of the project is to have all the vine images from the scans compile in an online dashboard. This dashboard will allow for navigation of the customers maps down to the block that has been imagined with the Flash system and ultimately down to the vine level. The gathering of images is designed to allow for you to see the same plant develop over growing season and allow for you to see the same plant year over year for the scans that are available. The development of the dashboard has gone from initial design to prototype and is now close to a 1.0 version for early customer usage. The newest customer facing version that includes the data from the scans under this program is almost ready to be released on to the website dashboard.bloomfield.ai and appointments will be made with all participants to review the results.

Technical Progress Made:

Bloomfield has made huge technical leaps forward over the past year, with advancements across hardware, software, AI, and data science.

- Flash evolved from version 1.0 to version 1.5, with increased data storage and an overhaul of the electrical systems for increased reliability and reduced power consumption.
- A largely automated data pipeline enables significantly faster turn-around times. Bloomfield is rapidly approaching their short-term goal of 48 to 72 turn-around time for delivering analysis to each customer through the dashboard.
- A scalable database was designed and deployed for storing the enormous amounts of data generated by the camera and the analysis processes.
- Various APIs were designed and deployed to enable structured interaction with the database by AI engineers, data scientists, and the web dashboard.
- Human labeling has changed from simple bounding boxes to pixel labeling, enabling the development of much more accurate AI.
- AI development practices have been completed overhauled to take advantage of the large volume of high-quality training data, API-based interaction with the database, and instant visualization of AI outputs through the web dashboard. New AI development time has been reduced from weeks to days.
- Detection of grapes with the ability to grade the color of each berry based upon a color scale and an early version of a heat map to evaluate results across a block.



Problems Encountered:

As partially described above, numerous challenges had to be overcome to enable a commercially successful scan program.

1. Incredibly long turn-around times were the result of inefficient academic research scripts and laborious, human-managed data processing steps.
2. Development of crop detection models were delayed during the growing season while Bloomfield was focused on other priorities. This accounted for much of the analysis to occur after the season had concluded.
3. Different grape varieties in different growth phases presents a significant AI development challenge. Bloomfield continues to work on the development of

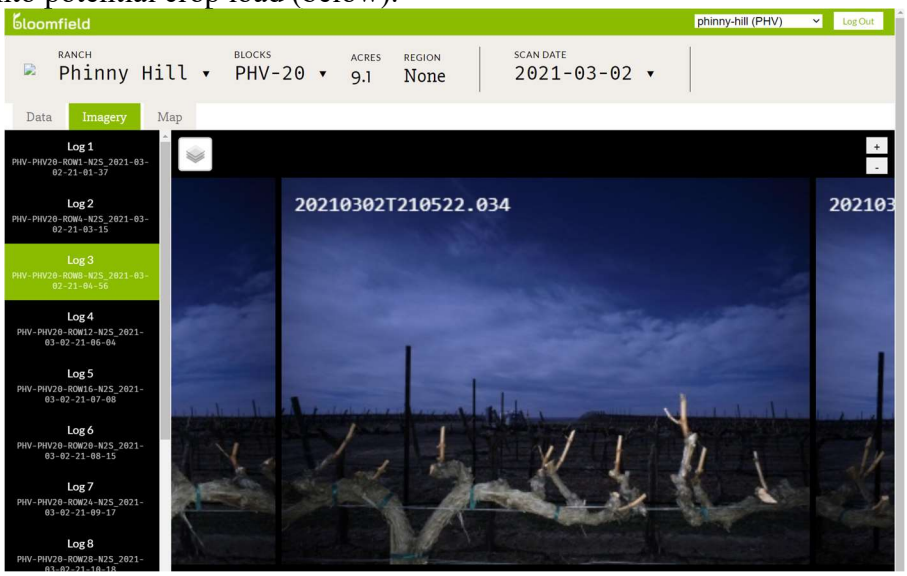
accurate and efficient AI models that can understand these differences and plans to begin unifying these models into a larger "superintelligence" over the course of 2021 and 2022.

4. The visibility into the grape canopy presents significant challenges to this technology and the method of capturing the data. Additionally, the visibility changes mainly on a grower-by-grower basis. This is something that is very visible while scanning blocks of the same variety in the same growing region all within a 10-mile radius. It could be that some method of heavy leaf removal needs to be applied to the process prior to scanning. Certainly, the most productive images come from scans conducted following leaf removal activities.
5. Bloomfield resource constraints and small team size severely limited the pace of development during 2020. As fund-raising efforts have begun to be more successful in 2021 (as we show the results of difficult infrastructure development and lessons learned), the Bloomfield team has been able to expand and increase their operational tempo.

Further Goals/Objectives:

Collab Wine Company in association with Bloomfield Robotics is currently working with a small subset of Grower Cooperators to continue scans during the 2021 growing season to further capture images and associated phenotypes to be analyzed along with testing what can be learned by imaging the same blocks across multiple years. The images will server to further advance the insights that can be achieved through the AI models.

Additionally, we are working on dormant season scans of blocks both prior to pruning and after with the thought that spurs and bud counts across an entire block could provide insights into potential crop load (below).



Along with models that detect clusters and inflorescence, the focus on inflorescence allows for clearer identification of the clusters along with a distinctive look (below).

The goal for the 2021 growing season would be to provide insights/information that can be inserted into the grower's own crop estimation formula while continuing to develop our own algorithms.

Discoveries from the Project:

Our primary discoveries revolve around systems architecture and the development of scalable cloud data processing practices. Delivering results for multiple growers in useful turn-around times has required evolution from academic research to commercial-grade infrastructure. In addition, various new techniques have been developed around view fusion and 3D reconstruction. Translating the stereo 2D imagery produced by the camera at 5 frames per second into accurate spatial reconstructions that can enable accurate counting is an area of ongoing computer science research, and we have made strides forward with new mosaic and point cloud techniques.

Intellectual Property applications filed:

Currently, there has not been any application for copyrights, trademarks, or patents associated with the work conducted in the Washington state vineyards associated with the research conducted during the 2020 growing season.

Conclusion:

Ultimately, the results of the project are disappointing, we set out to be able to determine the final crop load in a grape block across a much greater plant population than can be determine by current estimation practices. We did not achieve this nor were we able to deliver to the growers' actionable information during the growing season, but in the efforts to achieve these lofty goals a lot of progress has been made on Bloomfield Flash product, procedures, and process. They have gone from not having the necessary AI models and pipeline to process the data, to being able to scan a vineyard in Washington state and have the images posted to the dashboard in under 5 days with a goal in improve

that down to 3 days. Bloomfield now also possess the necessary images, expertise, and process to take an existing AI model add additional training and post the results to a consumer facing web portal. These abilities did not exist at the start of this project. With the involvement of the Washington Wine Commission, Collab Wine Company has played a significant role in collecting images specific to Washington vineyards, offering Bloomfield guidance and expertise that will help advance the Bloomfield product while making it useful to Washington growers and wineries. It is my anticipation that we will start to see information that can be used in viticulturist yield estimation formulas during the 2021 growing season with Bloomfield's own yield estimation calculations showing up for the 2022 vintage. I would like to thank the Washington Wine Commission and the research committee for committing resources to this project and allowing the advancement of the product based upon Washington State vineyards and practices. This should serve the industry well as Bloomfield continues to develop.